

Revolutionizing Statistical Programming with AI/ML – A Focus ADaM Programming

Francis Kendall,
Executive Director,
Statistical Programming,
Digitaland Data Sciences,
Biogen

June 2025





Changes in the Workplace





Evolution of Listening to Music

- What changed ?
- What didn't change?

Today's AI Environment at work



Email Filtering - AI helps filter out spam and prioritize important emails in our inboxes.



Data Analysis - AI tools analyze large datasets to uncover trends, patterns, and insights that inform business decisions.



Human Resources - AI streamlines recruitment by screening resumes, scheduling interviews, and even conducting initial



Cybersecurity - AI detects and responds to security threats in real-time, protecting sensitive company data.



Virtual Meetings - AI enhances virtual meetings with features like real-time transcription, translation, and meeting summaries.



Document Automation - AI automates the creation, editing, and management of documents, reducing manual effort and errors.



Personal Assistants - AI-powered personal assistants like Microsoft 365 Copilot help schedule meetings, set reminders, and manage daily



Statistical Programming Key Responsibilities and Tasks in this New Era

Data Management & Wrangling

Statistical programmers are now responsible for managing large datasets, ensuring data integrity, and maintaining accurate records. Data wrangling involves cleaning and transforming raw data into a usable format, ensuring data quality for analysis.

Importance of Data Analytics

A strong understanding of data analytics methodologies is essential for effective evidence generation and decision-making.

Interdisciplinary Collaboration

Collaboration among data scientists, statisticians, and domain experts enhances the quality of evidence generated from data.

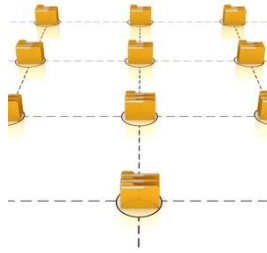
Robust Analytical Framework

Integrating data science into evidence generation enhances the analytical framework, allowing for deeper insights.

Machine Learning Insights

Machine learning is utilized to extract actionable insights that drive informed decision-making processes.

Data Accessibility and Inconsistent Standards



Challenge of Data Accessibility

Ensuring data accessibility across various departments is crucial for effective communication and collaboration.



Importance of Consistent Standards

Maintaining consistent data standards is essential for reliable data analysis and decision-making.



Impact on Collaboration

Inconsistent data practices can hinder effective collaboration and lead to misunderstandings among teams.



Importance of Data Governance

Organizations must implement strong data governance practices to manage and protect data, ensuring it is accessible to all stakeholders.

The Role of ADaM Programming

Why ADaM Matters

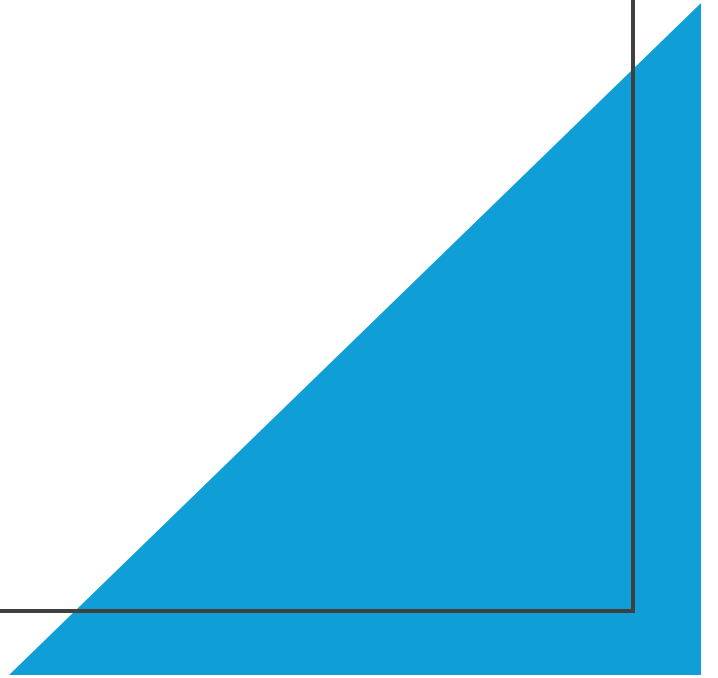
- Foundation for statistical analysis
- Critical for regulatory submissions
- Must comply with CDISC standards

Challenges in ADaM Generation

- Manual programming errors
- Time-consuming processes and validation
- Difficulty maintaining traceability and compliance especially across studies
- Data Quality Issues
- Regulatory Complexity
- High risk of human error
- Need for Skilled Personnel



Biogen Example



BIOGEN - Auto-ADaM Overview

Functionality / Purpose

- A web-based interactive application that is part of an integrated suite of tools
- Augments ADaM specification metadata captured in another tool to support standardized programmatic derivations
- Presents consolidated information to end-user pulled from various Biogen sources, including Biogen standards so the programmer has one-stop-shop for info
- Leverage SAS macros to enable more standard approach to analysis variable derivations, eg, deriving durations between two time-points use same macro across all domains
- Accommodate user-defined “custom code” for limited set of more complex variables

Biogen Auto-ADaM Overview

Benefits

- Efficiency: Reduce time and effort to produce working ADaM datasets
- Standards: Ensure greater alignment with Biogen Standards
- Standards: Adopting a more standard program structure leading to greater program similarity across studies/compounds
- Quality, Efficiency: Provides an environment for standard run-time “edit-checks” of derivation input/output, produce standard “edit-check” report providing “instant” feedback to programmer to assess program execution correctness
- Forward-Looking: Accumulate actual ADxx metadata allowing for retrospective analysis of compliance / trends across studies / compounds to inform process adjustments

Key Themes to Apply AI for ADaM

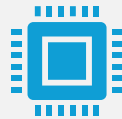
Automated ADaM Dataset Creation



Natural Language Processing (NLP): AI can interpret SAPs (Statistical Analysis Plans) and metadata to auto-generate ADaM specifications.



Template Recognition: ML models learn from historical datasets to apply consistent structures and naming conventions.



Code Generation: AI tools can write reusable SAS or Python code to generate ADaM datasets from SDTM inputs.



Benefit: Reduces manual programming time and ensures consistency across studies.

Intelligent Validation



Rule-Based & ML Validation Engines: AI can apply CDISC rules and learn from past validation errors to flag issues.



Anomaly Detection: ML models identify outliers or inconsistencies in derived variables or population flags.



Automated QC: AI can compare independently generated datasets for discrepancies, reducing double programming efforts.



Benefit: Enhances data quality and reduces human error.

Enhanced Traceability



Lineage Mapping: AI tools can automatically map the lineage from raw data → SDTM → ADaM → TLFs (Tables, Listings, Figures).



Interactive Traceability Reports: AI can generate visual traceability matrices and audit trails.



Metadata Linking: AI links variables to their source and transformation logic, improving transparency.



Benefit: Improves regulatory compliance and audit readiness.

Future Outlook



Agentic AI: Self-directed agents that can manage entire workflows—from reading protocols to generating submission-ready datasets.



Integration with Cloud Platforms: Real-time collaboration and validation across global teams.

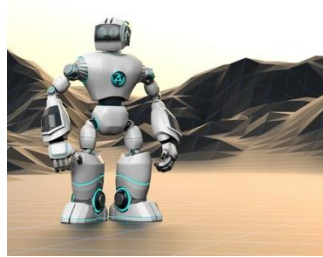


Regulatory Acceptance: AI-generated datasets are increasingly being reviewed and accepted by regulatory bodies with proper documentation.

AI Agents



AI Agents Abilities



Autonomous Entities

AI agents operate independently, making decisions based on their perception of the environment and pre-defined goals.



Data Processing

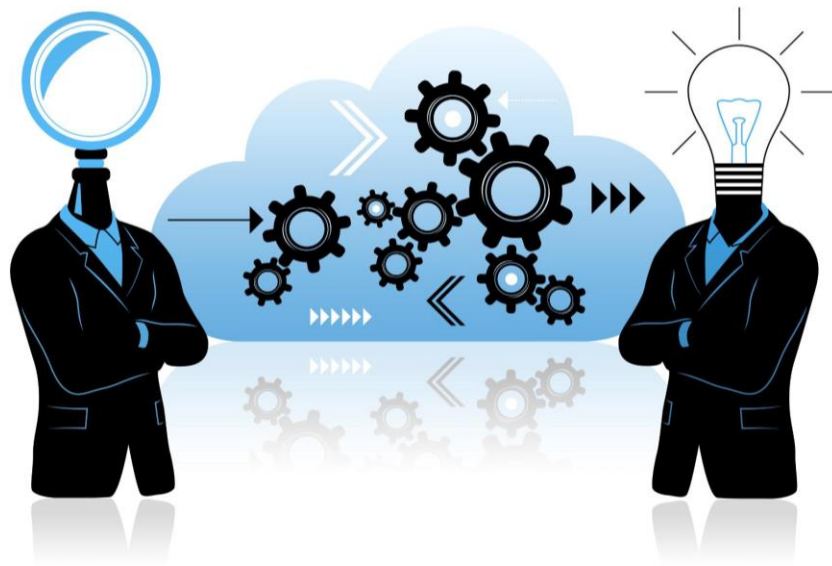
These agents utilize algorithms to process data, enabling them to learn and adapt their actions based on experiences.



Decision Making

AI agents make informed decisions to take actions that align with their specific goals, enhancing their effectiveness.

Design and Functionality of AI Agents



Automation of Repetitive Tasks

AI agents can automate repetitive tasks, freeing up time for developers to focus on more strategic initiatives.

Data Analysis Capabilities

These AI agents have advanced data analysis capabilities, helping to extract insights and make data-driven decisions.

Regulatory Compliance

AI agents ensure compliance with regulatory standards, reducing the risk of human error and improving accuracy in programming.

Enhanced Programming Experience

The ultimate goal of AI agents is to enhance the overall programming experience by streamlining workflows and improving efficiency.

AI Agent Use Cases in ADaM Programming

Automated Dataset Generation

Creates ADaM datasets from SDTM using metadata and AI rules.

Predictive Error Detection

Identifies potential errors using historical data and ML.

Real-Time Compliance Validation

Ensures datasets meet CDISC standards during creation.

Traceability Mapping

Generates traceability between SDTM and ADaM datasets.

Documentation Support

Drafts reviewer guides and define.xml files automatically.

Adaptive Learning

Learns from feedback to improve dataset generation and validation.

Risks and Challenges



GXP AND
AUDITABILITY



BIAS AND
OVERFITTING



INTERPRETABILITY
ISSUES



REGULATORY
ACCEPTANCE

Conclusion



Benefits: efficiency, accuracy, scalability



Call to action: embracing AI in biometrics workflows



Q&A
